Introduction aux méthodes Monte Carlo

Babahenini Med Chaouki

2015/2016

Expérience aléatoire et probabilité

- Suite d’évènements dont les variables varient dans des intervalles de valeurs connus, sans que les valeurs précises le soient avec une certitude absolue
- La fréquence d’occurrence de chaque valeur est utilisée comme mesure relative de sa certitude, ou probabilité d’occurrence

$$P(x=\alpha)=\frac{nombre\,de\;fois\,que\,l'on\,observe\;\;x=\alpha}{nombre\,d'observations\,des\,valeurs\,de\,x}$$

Approche évènementielle

- Expérience aléatoire : Suite d'évènements pris dans un espace Ω sur lequel est définie une probabilité P.
  - Un événement est une partie de Ω notée A.
  - Chaque événement possède une probabilité P(A) telle que, dans l'espace probabilisé (Ω, A, P), on a :

    $P(\Omega)=1$ (il est certain qu'un évènement se produise, sans savoir lequel)

    $P(A) + P(A^c)=1$ (Si un évènement ne se produit pas, un des autres se produira)
  - La loi de probabilité triviale est :

    Cas discret

    $$P(A) = \frac{Card\ (A)}{Card\ (\Omega)}$$

    Cas continu

    $$P(A)=\int_A p(x)dx$$

Approche par variable aléatoire

- Variable numérique dont les valeurs dépendent des résultats d'une expérience aléatoire
  - Représente une projection de (Ω, A, P) dans un espace numérique
  - Permet le calcul des probabilités par des méthodes de l'analyse mathématique au lieu de raisonner sur des ensembles
- Variable aléatoire discrète : ses valeurs sont dénombrables
  - Espérance : la valeur moyenne

$$E(X) = \sum_E x\, P(X = x)$$

  - Variance : l'écart quadratique moyen par rapport à la valeur moyenne

$$V(X) = E\left(\left[X - E(X)\right]^2\right)$$

Probabilités continues

- Une variable aléatoire décrit les résultats possibles d'une expérience.
- Associé à une variable aléatoire y, F(y) est une fonction de distribution de probabilité. Cette fonction donne la probabilité avec laquelle un événement se produit avec un résultat inférieur ou égal à la valeur de y.
- F (y) = probabilité de la variable aléatoire ≤ y
- F(y) est une fonction non décroissante, et est non-négatif sur l'ensemble du domaine de la variable aléatoire.
- Soit une PDF, on construit la fonction de distribution de probabilité correspondante par: $F(y) = \int_{-\infty}^{y} f(x)dx$

Probabilités continues

- La distribution des valeurs que prend x est décrite donc par une fonction de distribution de probabilité p, f(x) (PDF) tel que:
  - On dit que x est distribué selon p, ou *x* ~ *p*
  - $p(x) \geq 0$
  - $$\int_{-\infty}^{\infty} p(x)dx = 1$$
- La probabilité pour que x prend une valeur dans l'intervalle [a, b] est: $$\mathrm{P}(x \in [a,b]) = \int_a^b p(x)dx$$

L’espérance mathématique

- En théorie des probabilités, l'**espérance mathématique** d'une variable aléatoire réelle est, intuitivement, la valeur que l'on s'attend à trouver, en moyenne, si l'on répète un grand nombre de fois la même expérience aléatoire.
- La valeur attendue de x ~ p est défini comme:

$$E(x) = \int x p(x) dx$$

- En tant que fonction d'une variable aléatoire est lui-même une variable aléatoire, la valeur attendue de f (x) est:

$$E(f(x)) = \int f(x) p(x) dx$$

- La valeur attendue d'une somme de variables aléatoires est la somme des valeurs attendues:

$$\mathrm{E}(x + y) = \mathrm{E}(x) + \mathrm{E}(y)$$

Variables aléatoires multi-dimensionnelles

- Pour un espace S, nous pouvons définir une pdf

$$p: S \rightarrow \Re$$

- Si x est une variable aléatoire, et x ~ p, la probabilité que x prend une valeur dans $S_i \subset S$ est :

$$\mathrm{P}(x \in S_i) = \int_{S_i} p(x) d\mu$$

Valeur attendue d'une fonction réelle $f: S \rightarrow \Re$ étend naturellement au cas multidimensionnel:

$$E(f(x)) = \int_S f(x) p(x) d\mu$$

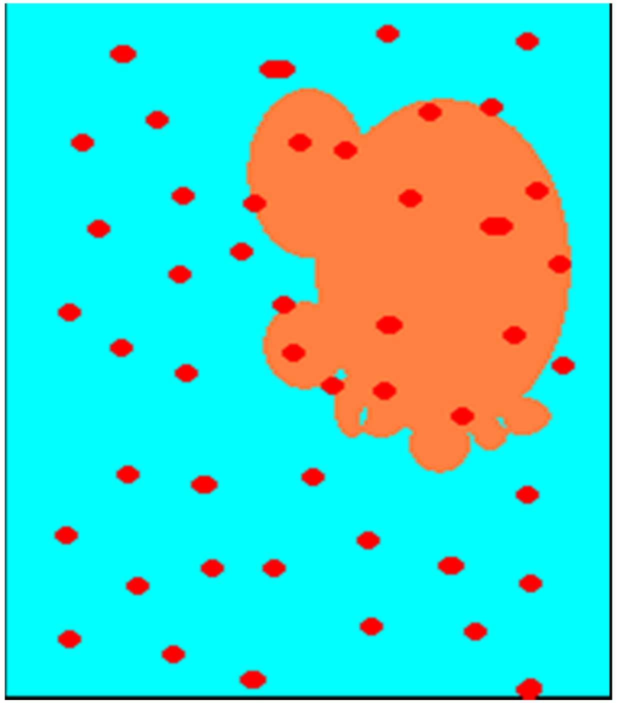
Variables aléatoires multi-dimensionnelles

- Dans notre cas S est souvent une zone ($d\mu = dA = dx.dy$) ou un ensemble de directions (points d'une sphère unité: $d\mu = d\omega = \sin\theta . d\theta . d\varphi$
- Une v.a. à 2 dimensions α est uniformément distribuée sur un disque de rayon R.

$$p(\alpha) = \frac{1}{\pi . R^2} \Rightarrow \mathrm{Proba}(\alpha \in S_i) = \int_{S_i} \frac{1}{\pi . R^2} . dA$$

Exemple introductif: Approximation de Pi par la méthode de Monte Carlo

- <u>Idée générale:</u> On place des points au hasard dans un domaine d'aire connu. Lorsque le nombre de points placés tend vers l'infini, la proportion des points « tombés » dans un sous domaine permet d'obtenir son aire.



- En pratique, il faut cependant pouvoir:
  - Placer des points aléatoirement dans le domaine
  - Compter ceux ayant atterri sur le sous domaine grâce à une formule.

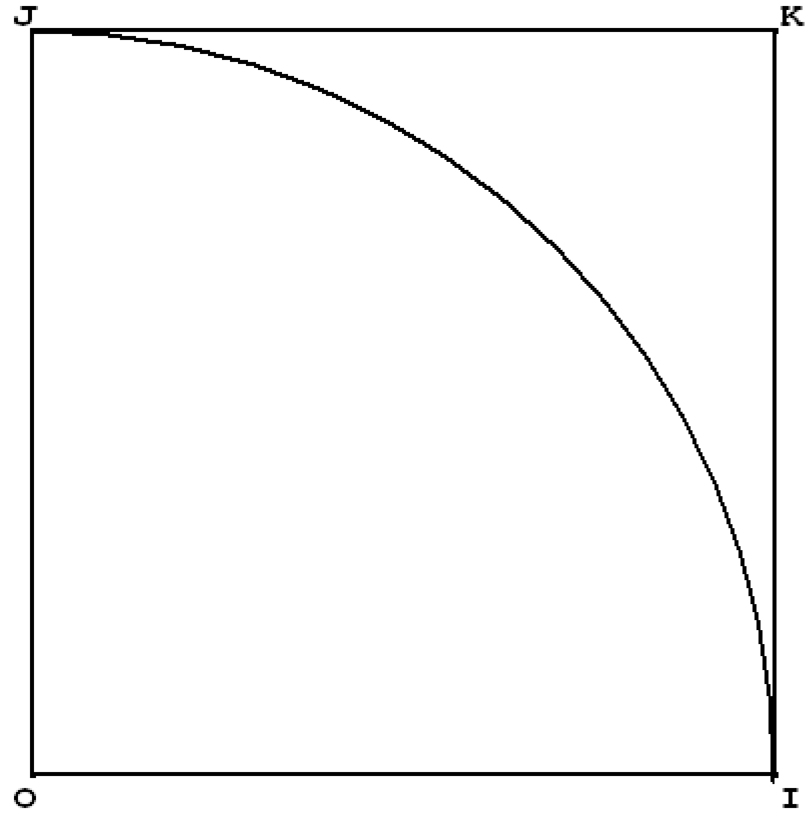
En pratique, il faut cependant pouvoir

- Placer des points aléatoirement dans le domaine
- Compter ceux ayant atterri sur le sous domaine grâce à une formule.

Situation étudiée:

- Dans le repère ortho-normal (O;I;J) l'aire du carré OIKJ vaut 1.

- On va utiliser la méthode de MC pour approcher l'aire du quart de disque c'est-à-dire de Pi/4.



Algorithme en langage naturel

VARIABLE

disque: compte le nombre de points situés à l'intérieur du quart de disque
n: nombre de points placés aléatoirement
x :abscisse d'un point au hasard dans le carré
y :ordonnée d'un point au hasard dans le carré

TRAITEMENT

Pour i de 1 à n

x prend une valeur aléatoire dans [0 ;1]

y prend une valeur aléatoire dans [0 ;1]

Si $x^2+y^2<1$ alors disque prend la valeur **disque** +1

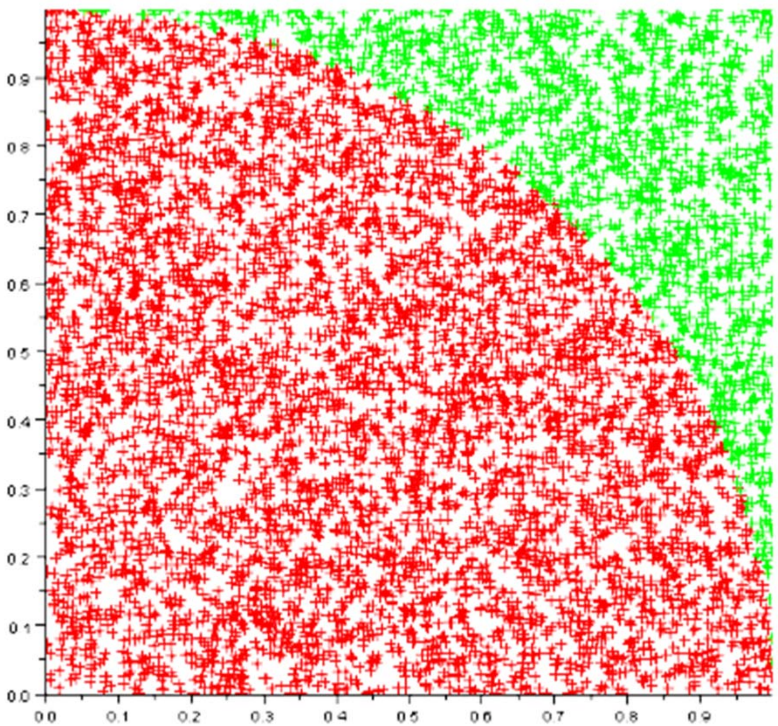
SORTIE

Afficher (disque /n)

Fichier Edition Rechercher Exécuter Débug Lan

```
clf;
disque=0;
for i=1:10^8
  x=rand();
  y=rand();
if x^2+y^2<1 then disque=disque+1;
                  plot(x,y,"+r")
  else plot(x,y,"+g")
end
end
afficher(4*disque/10^8)
```

Visualisation des résultats




```
-1->clf;

-1->disque=0;

-1->for i=1:10^8
-1->   x=rand();
-1->   y=rand();
-1->if x^2+y^2<1 then disque=disque+1;
-1->
-1->end
-1->end
-1->afficher(4*disque/10^8)

    3.1416378
```

Pourquoi ça marche ?

- C’est la loi des grands nombres

- On pose $p = \dfrac{\text{aire du quart de disque}}{\text{aire du carré}}$

- On pose f_{obs} la fréquence observée de points situés dans le quart de disque.
- Alors en plaçant n points aléatoirement, si n est assez grand,

on a plus de 95% de chances que p soit compris dans

$$\left[f_{obs} - \frac{1}{\sqrt{n}} ; f_{obs} + \frac{1}{\sqrt{n}} \right]$$

Points forts/Points faibles de la méthode

Points forts

- C’est simple!
- Ca marche !

Point faible

- La méthode est très coûteuse en calcul à cause de la lenteur de la convergence:

  pour avoir n décimales, il faut placer 10^(2n) points !

  Du coup ici, on ne voit pas beaucoup de décimales de Pi...mais on voit facilement 3,14

Intégration Monte Carlo

- Supposons que nous avons une fonction f (x) définie sur le domaine x ϵ [a, b].
  Nous voulons évaluer l'intégrale:
  $$I = \int_a^b f(x)dx$$
- La méthode de Monte Carlo considère N échantillons choisis au hasard avec pdf p (x), pour estimer l'intégrale.
- Nous obtenons l'estimateur suivant:

$$\langle I \rangle = \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{p(x_i)}$$

Intégration Monte Carlo

- Trouver la valeur estimée de l'estimateur, on obtient:

$$E[\langle I \rangle] = E\left[\frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)}\right] = \frac{1}{N}\sum_{i=1}^{N}E\left[\frac{f(x_i)}{p(x_i)}\right] = \frac{1}{N}N\int\frac{f(x)}{p(x)}p(x)dx$$

$$= \int f(x)dx = I$$

- En d'autres termes:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)} = I$$

Variance

- La **variance** est une mesure servant à caractériser la dispersion d'un échantillon ou d'une distribution. Elle indique de quelle manière la variable aléatoire se disperse autour de sa moyenne ou son espérance.
  - Une variance de zéro signale que toutes les valeurs sont identiques.
  - Une petite variance est signe que les valeurs sont proches les unes des autres
  - alors qu'une variance élevée est signe que celles-ci sont très écartées.

$$\sigma^2[x] = \int_a^b (x - E[x])^2 . p(x).dx = E[x^2] - E[x]^2$$

Variance

- La variance de l'estimateur est:

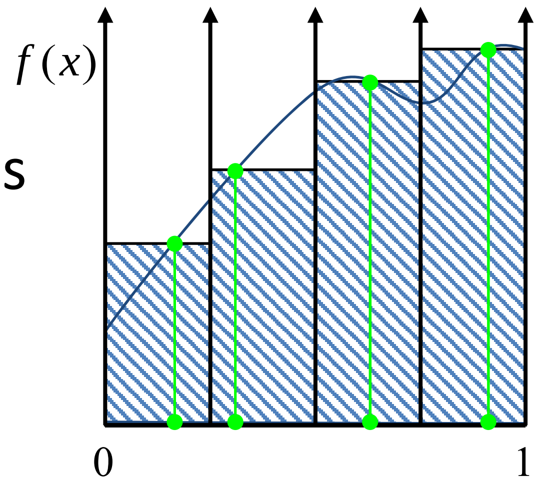
$$\sigma^2 = \frac{1}{N}\int\left(\frac{f(x)}{p(x)} - I\right)^2 p(x)dx$$

- Comme l'erreur de l'estimation est proportionnelle à σ, l'erreur est proportionnelle à $\frac{1}{\sqrt{N}}$
  - Donc, pour réduire de moitié l'erreur, nous devons utiliser quatre fois plus d'échantillons.

Réduction de la variance

- Pour avoir un bon estimateur, il faut :
  - Le plus d'échantillons possible.
  - f/p doit avoir une petite variance doivent avoir des parties similaires.
- **Echantillonnage d'importance:** parce que si p est grande lorsque f est grande et petite lorsque f est petit, il y aura plus l'échantillon dans les régions importantes.
- Partitionner le domaine de l'intégrale dans plusieurs régions plus petites et évaluer l'intégrale en tant que la somme des intégrales sur les régions plus petites Cela se appelle **l'échantillonnage stratifié**

Réduction de la variance: échantillonnage stratifié



- L'espace D est divisé en sous espaces disjoints
- $D_i=[x_i,x_{i+1}]$: sous espace $\sum_i P_i=1$

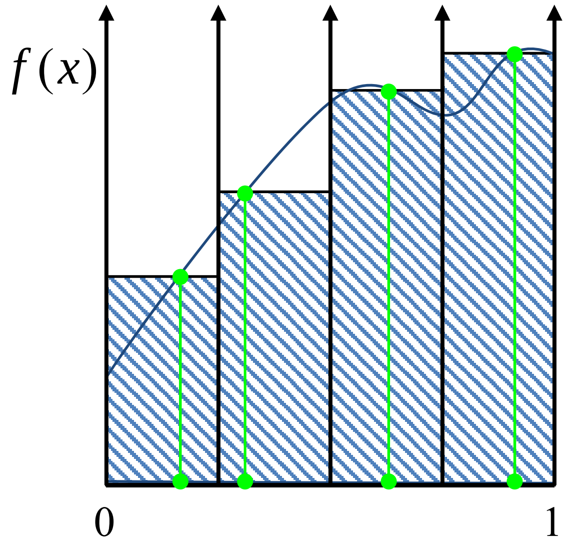
$$I=\int_0^1 f(x)dx=\int_0^{x_1}(f(x)/p(x))p(x)dx+$$

$$\int_{x1}^{x2}(f(x)/p(x))p(x)dx+$$

$$\ldots+\int_{x_{n-1}}^{1}(f(x)/p(x))p(x)dx$$

Réduction de la variance: échantillonnage stratifié

- N_i échantillons du sous espace i
- n sous espaces
- $p(x)/P_i$: pdf pour un sous espace i



$$\int_{Di} p(x)dx = Pi$$

$$\int_{Di} (p(x)/P_i)dx = 1$$

$$I = \sum_{i=0}^{n-1} Pi \int_{Di} \frac{f(x)}{p(x)} . \frac{p(x)}{P_i} dx$$

$$I \approx \sum_{i=0}^{i=n-1} \frac{Pi}{Ni} \sum_{ki=0}^{Ni} \frac{f(X_{ki})}{p(X_{ki})}$$

Echantillonnage de variables aléatoires

- Etant donné une pdf $p(x)$, définie sur un intervalle $[x_{min}, x_{max}]$, nous pouvons échantillonner la va $x \sim p$ à partir d'un ensemble nombres aléatoires uniformes $\xi_i \in [0,1]$ $$prob(X < x) = P(x) = \int_{x_{\min}}^{x} p(\mu)d\mu$$

- Pour cela, nous avons besoin de la CDF, *cumulative probability distribution function*:
  - Pour avoir x_i nous transformons ξ_i: $x_i = P^{-1}(\xi_i)$
  - P^{-1} existe toujours pour toutes les pdfs valides

Exemple

- Echantillonnage de la pdf $p(x)=3x^2/2$, $x \in [-1, 1]$:
  - Premièrement nous avons besoin de trouver P(x):

$$P(x) = \int 3x^2 / 2dx = x^3 / 2 + C$$

  - On choisit C tel que P(-1)=0 et P(1)=1, nous trouvons $P(x)=(x^3+1)/2$ et $P^{-1}(\xi)=$ $\sqrt[3]{2\xi - 1}$

  - Donc, pour tout nombre aléatoire $\xi_i \in [0,1]$,

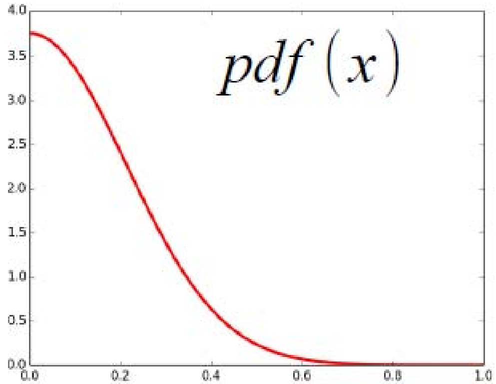
$$x_i = \sqrt[3]{2\xi_i - 1}$$

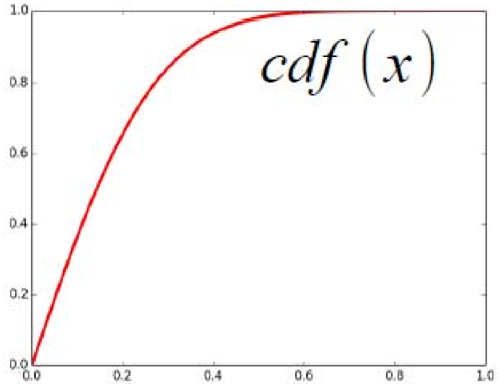
Échantillonnage selon une densité

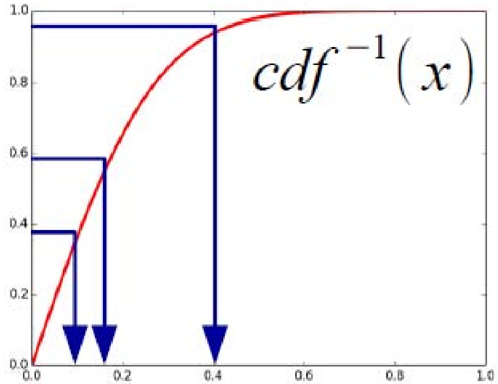
- Nous disposons seulement d'échantillonnages uniformes...
- Nous devons construire un autre échantillonnage, *pdf* donnée
- Méthodologie :
  - Déterminer la fonction de répartition *cdf* (Cumulative Density Function)

$$cdf(x) = \int_0^x pdf(t)\,dt$$

  - Inverser la *cdf* , pour finalement obtenir : $u_i' = cdf^{-1}(u_i)$
  - Echantillonnage d'importance : $val = cdf^{-1}(\xi)$



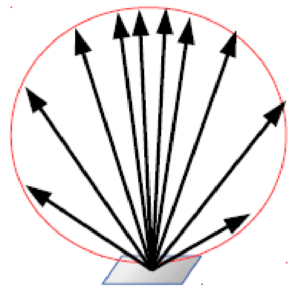




Exemple1: Modèle de Lambert

$$pdf(m) \approx \cos\theta_m, \quad avec \quad \int_\Omega \cos\theta_m \, d\omega_m \stackrel{?}{=} 1 \quad / \quad d\omega_m = \sin\theta_m \, d\theta_m \, d\varphi_m$$

Modèle de Lambert : $L_o = L_i \times \frac{K_d}{\pi} \cos\theta_i$

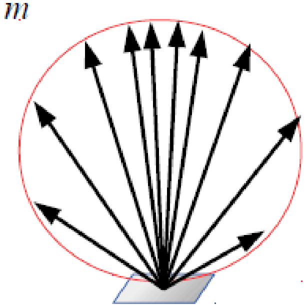


$$pdf(m) \approx \cos\theta_m, \quad avec \quad \int_\Omega \cos\theta_m \, d\omega_m \stackrel{?}{=} 1 \quad / \quad d\omega_m = \sin\theta_m \, d\theta_m \, d\varphi_m$$

$$\int_{\varphi_m=0}^{2\pi} \int_{\theta_m=0}^{\pi/2} \cos\theta_m \sin\theta_m \, d\theta_m \, d\varphi_m$$

$$2\pi \int_{\theta_m=0}^{\pi/2} \cos\theta_m \sin\theta_m \, d\theta_m$$

$$2\pi \left[- \frac{1}{2} \cos^2\theta_m \right]_0^{\pi/2} = \pi$$



Exemple 1: Modèle de Lambert

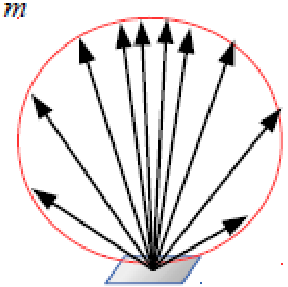
- $pdf(m) \approx \cos\theta_m, \quad avec \int_\Omega \cos\theta_m \, d\omega_m \overset{?}{=} 1 \quad / \quad d\omega_m = \sin\theta_m \, d\theta_m \, d\varphi_m$

$$\int_{\varphi_m=0}^{2\pi} \int_{\theta_m=0}^{\pi/2} \cos\theta_m \sin\theta_m \, d\theta_m \, d\varphi_m$$

$$2\pi \int_{\theta_m=0}^{\pi/2} \cos\theta_m \sin\theta_m \, d\theta_m$$

$$2\pi \left[-\frac{1}{2} \cos^2\theta_m \right]_0^{\pi/2} = \pi$$

$$\boxed{pdf(m) = \frac{1}{\pi} \cos\theta_m}$$



- $cdf(x) = \int_0^x pdf(m) \, dm = \frac{1}{\pi} \int_{\varphi=0}^{2\pi} \int_{\theta_m=0}^{x} \cos\theta_m \sin\theta_m \, d\theta_m$

$$= \frac{1}{\pi} 2\pi \int_{\theta_m=0}^{\pi/2} \cos\theta_m \sin\theta_m \, d\theta_m = \left[-\cos^2\theta_m \right]_0^x = \boxed{1 - \cos^2 x}$$

- $cdf^{-1}(x) = arcos(\sqrt{1-x})$

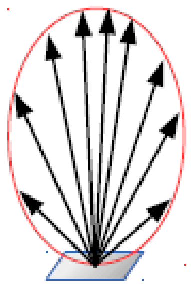
$$\boxed{\begin{array}{l} \varphi_m^i = \xi_1 2\pi \\ \theta_m^i = arcos(\sqrt{1-\xi_2}) \end{array}}$$

Exemple 2: Modèle de Phong modifié

- Echantillonnage lobe perpendiculaire : $pdf(m)=\frac{n+2}{2\pi}\cos^n\theta_m\cos\theta_m$

$$cdf(m)=\int_{\theta=0}^{x}(n+2)\cos^{n+1}\theta_m\sin\theta_m\,d\theta_m=[-\cos^{n+2}\theta_m]_0^x=\boxed{1-\cos^{(n+2)}x}$$

- Choix par importance : $\boxed{\begin{array}{l}\varphi_m^i=\xi_1 2\pi\\ \theta_m^i=arcos\left((1-\xi_2)^{\frac{1}{n+2}}\right)\end{array}}$



- Une fois l'échantillon choisi :
  - ✔ *Construire le vecteur « miroir » par rapport à la direction d'observation*
  - ✔ *Reporter (rotation) la direction échantillonnée*